

SEQUENCE LISTING PART OF THE DESCRIPTION

SEQ. ID. NO. 1 - Wild type gagpol sequence for strain HXB2 (accession no. K03455)

ATGGGTGCGA	GAGCGTCAGT	ATTAAGCGGG	GGAGAATTAG	ATCGATGGGA	AAAAATTCGG	60
TTAAGGCCAG	GGGGAAGAA	AAAATATAAA	TTAAAACATA	TAGTATGGGC	AAGCAGGGAG	120
CTAGAACGAT	TCGCAGTTAA	TCCTGGCCTG	TTAGAAACAT	CAGAAGGCTG	TAGACAAATA	180
CTGGGACAGC	TACAACCATC	CCTTCAGACA	GGATCAGAAG	AACTTAGATC	ATTATATAAT	240
ACAGTAGCAA	CCCTCTATTG	TGTGCATCAA	AGGATAGAGA	TAAAAGACAC	CAAGGAAGCT	300
TTAGACAAGA	TAGAGGAAGA	GCAAAACAAA	AGTAAGAAAA	AAGCACAGCA	AGCAGCAGCT	360
GACACAGGAC	ACAGCAATCA	GGTCAGCCAA	AATTACCCTA	TAGTGCAGAA	CATCCAGGGG	420
CAAATGGTAC	ATCAGGCCAT	ATCACCTAGA	ACTTTAAATG	CATGGGTAAA	AGTAGTAGAA	480
GAGAAGGCTT	TCAGCCCAGA	AGTGATACCC	ATGTTTTTCAG	CATTATCAGA	AGGAGCCACC	540
CCACAAGATT	TAAACACCAT	GCTAAACACA	GTGGGGGGAC	ATCAAGCAGC	CATGCAAATG	600
TTAAAAGAGA	CCATCAATGA	GGAAGCTGCA	GAATGGGATA	GAGTGCATCC	AGTGCATGCA	660
GGGCCTATTG	CACCAGGCCA	GATGAGAGAA	CCAAGGGGAA	GTGACATAGC	AGGAACCTACT	720
AGTACCCTTC	AGGAACAAAT	AGGATGGATG	ACAAATAATC	CACCTATCCC	AGTAGGAGAA	780
ATTTATAAAA	GATGGATAAT	CCTGGGATTA	AATAAAATAG	TAAGAATGTA	TAGCCCTACC	840
AGCATTCTGG	ACATAAGACA	AGGACCAAAG	GAACCCTTTA	GAGACTATGT	AGACCGGTTT	900
TATAAACTC	TAAGAGCCGA	GCAAGCTTCA	CAGGAGGTAA	AAAATTGGAT	GACAGAAACC	960
TTGTTGGTCC	AAAATGCGAA	CCCAGATTGT	AAGACTATTT	TAAAAGCATT	GGGACCAGCG	1020
GCTACACTAG	AAGAAATGAT	GACAGCATGT	CAGGGAGTAG	GAGGACCCGG	CCATAAGGCA	1080
AGAGTTTTGG	CTGAAGCAAT	GAGCCAAGTA	ACAAATTCAG	CTACCATAAT	GATGCAGAGA	1140
GGCAATTTTA	GGAACCAAAG	AAAGATTGTT	AAGTGTTC	ATTGTGGCAA	AGAAGGGCAC	1200
ACAGCCAGAA	ATTGCAGGGC	CCCTAGGAAA	AAGGGCTGTT	GGAAATGTGG	AAAGGAAGGA	1260
CACCAATGA	AAGATTGTAC	TGAGAGACAG	GCTAATTTTT	TAGGGAAGAT	CTGGCCTTCC	1320
TACAAGGGAA	GGCCAGGGAA	TTTTCTTCAG	AGCAGACCAG	AGCCAACAGC	CCCACCAGAA	1380
GAGAGCTTCA	GGTCTGGGGT	AGAGACAACA	ACTCCCCCTC	AGAAGCAGGA	GCCGATAGAC	1440
AAGGAAGTGT	ATCCTTTAAC	TTCCCTCAGG	TCACTCTTTG	GCAACGACCC	CTCGTCACAA	1500
TAAAGATAGG	GGGGCAACTA	AAGGAAGCTC	TATTAGATAC	AGGAGCAGAT	GATACAGTAT	1560
TAGAAGAAAT	GAGTTTGCCA	GGAAGATGGA	AACCAAAAAT	GATAGGGGGA	ATTGGAGGTT	1620
TTATCAAAGT	AAGACAGTAT	GATCAGATAC	TCATAGAAAT	CTGTGGACAT	AAAGCTATAG	1680
GTACAGTATT	AGTAGGACCT	ACACCTGTCA	ACATAATTGG	AAGAAATCTG	TTGACTCAGA	1740
TTGGTTGCAG	TTTAATTTT	CCCATTAGCC	CTATTGAGAC	TGTACCAGTA	AAATTAAGC	1800
CAGGAATGGA	TGGCCCAAAA	GTTAAACAAT	GGCCATTGAC	AGAAGAAAAA	AGAAGAGCAT	1860
TAGTAGAAAT	TTGTACAGAG	ATGGAAAAGG	AAGGGAAAAT	TTCAAAAATT	GGGCCTGAAA	1920
ATCCATACAA	TACTCCAGTA	TTTGCCATAA	AGAAAAAAGA	CAGTACTAAA	TGGAGAAAAT	1980
TAGTAGATTT	CAGAGAACCT	AATAAGAGAA	CTCAAGACTT	CTGGGAAGTT	CAATTAGGAA	2040
TACCACATCC	CGCAGGGTTA	AAAAAGAAAA	AATCAGTAAC	AGTACTGGAT	GTGGGTGATG	2100
CATATTTTTT	AGTTCCCTTA	GATGAAGACT	TCAGGAAGTA	TACTGCATTT	ACCATACCTA	2160
GTATAACAA	TGAGACACCA	GGGATTAGAT	ATCAGTACAA	TGTGCTTCCA	CAGGGATGGA	2220
AAGGATCACC	AGCAATATTC	CAAAGTAGCA	TGACAAAAAT	CTTAGAGCCT	TTTAGAAAAAC	2280
AAAATCCAGA	CATAGTTATC	TATCAATACA	TGGATGATTT	GTATGTAGGA	TCTGACTTAG	2340
AAATAGGGCA	GCATAGAACA	AAAATAGAGG	AGCTGAGACA	ACATCTGTTG	AGGTGGGGAC	2400
TTACCACACC	AGACAAAAAA	CATCAGAAAG	AACCTCCATT	CCTTTGGATG	GGTTATGAAC	2460
TCCATCCTGA	TAAATGGACA	GTACAGCCTA	TAGTGCTGCC	AGAAAAAGAC	AGCTGGACTG	2520
TCAATGACAT	ACAGAAAGTTA	GTGGGGAAAT	TGAATTGGGC	AAGTCAGATT	TACCCAGGGA	2580
TTAAAGTAAG	GCAATTATGT	AAACTCCTTA	GAGGAACCAA	AGCACTAACA	GAAGTAATAC	2640
CACTAACAGA	AGAAGCAGAG	CTAGAAGTGG	CAGAAAACAG	AGAGATTCTA	AAAGAACCAG	2700
TACATGGAGT	GTATTATGAC	CCATCAAAAAG	ACTTAATAGC	AGAAATACAG	AAGCAGGGGC	2760
AAGGCCAATG	GACATATCAA	ATTTATCAAG	AGCCATTTAA	AAATCTGAAA	ACAGGAAAAAT	2820
ATGCAAGAAT	GAGGGGTGCC	CACACTAATG	ATGTAAAAACA	ATTAACAGAG	CGAGTGCAAA	2880
AAATAACCAC	AGAAAGCATA	GTAATATGGG	GAAAGACTCC	TAAATTTAAA	CTGCCCATAC	2940
AAAAGGAAAC	ATGGGAAACA	TGGTGGACAG	AGTATTGGCA	AGCCACCTGG	ATTCCTGAGT	3000
GGGAGTTTGT	TAATACCCCT	CCCTTAGTGA	AATTATGGTA	CCAGTTAGAG	AAAGAACCCA	3060
TAGTAGGAGC	AGAAACCTTC	TATGTAGATG	GGGCAGCTAA	CAGGGAGACT	AAATTAGGAA	3120
AAGCAGGATA	TGTTACTAAT	AGAGGAAGAC	AAAAAGTTGT	CACCCTAACT	GACACAACAA	3180
ATCAGAAGAC	TGAGTTACAA	GCAATTTATC	TAGCTTTGGA	GGATTGCGGA	TTAGATCAGT	3240
ACATAGTAAC	AGACTCACAA	TATGCATTAG	GAATCATTCA	AGCACAACCA	GATCAAAGTG	3300
AATCAGAGTT	AGTCAATCAA	ATAATAGAGC	AGTTAATAAA	AAAGGAAAAG	GTCTATCTGG	3360
CATGGGTACC	AGCACACAAA	GGAATTGGAG	GAAATGAACA	AGTAGATAAA	TTAGTCAGTG	3420
CTGGAATCAG	GAAAGTACTA	TTTTTAGATG	GAATAGATAA	GGCCCAAGAT	GAACATGAGA	3480
AATATCACAG	TAATTGGAGA	GCAATGGCTA	GTGATTTTAA	CCTGCCACCT	GTAGTAGCAA	3540

09936572.121101

AAGAAATAGT	AGCCAGCTGT	GATAAATGTC	AGCTAAAAGG	AGAAGCCATG	CATGGACAAG	3600
TAGACTGTAG	TCCAGGAATA	TGGCAACTAG	ATTGTACACA	TTTAGAAGGA	AAAGTTATCC	3660
TGGTAGCAGT	TCATGTAGCC	AGTGGATATA	TAGAAGCAGA	AGTTATTCCA	GCAGAAACAG	3720
GGCAGGAAAC	AGCATATTTT	CTTTTAAAT	TAGCAGGAAG	ATGGCCAGTA	AAAACAATAC	3780
ATACTGACAA	TGGCAGCAAT	TTCACCGGTG	CTACGGTTAG	GGCCGCCTGT	TGGTGGGCGG	3840
GAATCAAGCA	GGAATTTGGA	ATTCCCTACA	ATCCCCAAG	TCAAGGAGTA	GTAGAATCTA	3900
TGAATAAAGA	ATTAAAGAAA	ATTATAGGAC	AGGTAAGAGA	TCAGGCTGAA	CATCTTAAGA	3960
CAGCAGTACA	AATGGCAGTA	TTCATCCACA	ATTTTAAAG	AAAAGGGGGG	ATTGGGGGGT	4020
ACAGTGCAGG	GGAAAGAATA	GTAGACATAA	TAGCAACAGA	CATACAAACT	AAAGAATTAC	4080
AAAAACAAAT	TACAAAATT	CAAAATTTTC	GGGTTTATTA	CAGGGACAGC	AGAAATTCAC	4140
TTTGAAAGG	ACCAGCAAAG	CTCCTCTGGA	AAGGTGAAGG	GGCAGTAGTA	ATACAAGATA	4200
ATAGTGACAT	AAAAGTAGTG	CCAAGAAGAA	AAGCAAAGAT	CATTAGGGAT	TATGGAAAAC	4260
AGATGGCAGG	TGATGATTGT	GTGGCAAGTA	GACAGGATGA	GGATTAG		4307

SEQ I.D. NO. 2 - gagpol-SYNgp - codon optimised gagpol sequence

ATGGGCGCCC	GCGCCAGCGT	GCTGTCGGGC	GGCGAGCTGG	ACCGCTGGGA	GAAGATCCGC	60
CTGCGCCCCG	GCGGCAAAAA	GAAGTACAAG	CTGAAGCACA	TCGTGTGGGC	CAGCCGCGAA	120
CTGGAGCGCT	TCGCCGTGAT	CCCCGGGCTC	CTGGAGACCA	GCGAGGGGTG	CCGCCAGATC	180
CTCGGCCAAC	TGCAGCCCAG	CCTGCAAAAC	GGCAGCGAGG	AGCTGCGCAG	CCTGTACAAC	240
ACCGTGGCCA	CGCTGTACTG	CGTCCACCAG	CGCATCGAAA	TCAAGGATAC	GAAAGAGGCC	300
CTGGATAAAA	TCGAAGAGGA	ACAGAATAAG	AGCAAAAAGA	AGGCCCAACA	GGCCGCCGCG	360
GACACCGGAC	ACAGCAACCA	GGTCAGCCAG	AACTACCCCA	TCGTGCAGAA	CATCCAGGGG	420
CAGATGGTGC	ACCAGGCCAT	CTCCCCCGC	ACGCTGAACG	CCTGGGTGAA	GGTGGTGGAA	480
GAGAAGGCTT	TTAGCCCGGA	GGTGATACCC	ATGTTCTCAG	CCCTGTGAGA	GGGAGCCACC	540
CCCCAAGATC	TGAACACCAT	GCTCAACACA	GTGGGGGGAC	ACCAGGCCG	CATGCAGATG	600
CTGAAGGAGA	CCATCAATGA	GGAGGCTGCC	GAATGGGATC	GTGTGCATCC	GGTGCACGCA	660
GGGCCCATCG	CACCGGGCCA	GATGCGTGAG	CCACGGGGCT	CAGACATCGC	CGGAACGACT	720
AGTACCCTTC	AGGAACAGAT	CGGCTGGATG	ACCAACAACC	CACCCATCCC	GGTGGGAGAA	780
ATCTACAAAC	GCTGGATCAT	CCTGGGCCTG	AACAAGATCG	TGCGCATGTA	TAGCCCTACC	840
AGCATCCTGG	ACATCCGCCA	AGGCCCGAAG	GAACCTTTTC	GCGACTACGT	GGACCGGTTG	900
TACAAAACGC	TCCGCGCCGA	GCAGGCTAGC	CAGGAGGTGA	AGAAGTGGAT	GACCGAACC	960
CTGCTGGTCC	AGAACGCGAA	CCCGGACTGC	AAGACGATCC	TGAAGGCCCT	GGGCCAGCG	1020
GCTACCCTAG	AGGAAATGAT	GACCGCCTGT	CAGGGAGTGG	GCGGACCCGG	CCACAAGGCA	1080
CGCGTCCCTG	CTGAGGCCAT	GAGCCAGGTG	ACCAACTCCG	CTACCATCAT	GATGCAGCGC	1140
GGCAACTTTC	GGAACCAACG	CAAGATCGTC	AAGTGCTTCA	ACTGTGGCAA	AGAAGGGCAC	1200
ACAGCCCGCA	ACTGCAGGGC	CCCTAGGAAA	AAGGTGCTGT	GGAAATGCGG	CAAGGAAGGC	1260
CACCAGATGA	AAGACTGTAC	TGAGAGACAG	GCTAATTTTT	TAGGGAAGAT	CTGGCCTTCC	1320
TACAAGGGAA	GGCCAGGGAA	TTTTCTTCAG	AGCAGACCAG	AGCCAACAGC	CCCACCAGAA	1380
GAGAGCTTCA	GGTCTGGGGT	AGAGACAACA	ACTCCCCCTC	AGAAGCAGGA	GCCGATAGAC	1440
AAGGAACTGT	ATCCTTTAAC	TTCCCTCAGA	TCACTCTTTG	GCAACGACCC	CTCGTCACAA	1500
TAAAGATAGG	GGGGCAGCTC	AAGGAGGCTC	TCCTGGACAC	CGGAGCAGAC	GACACCGTGC	1560
TGGAGGAGAT	GTCGTTGCCA	GGCCGCTGTA	AGCCGAAGAT	GATCGGGGGA	ATCGGCGGTT	1620
TCATCAAGGT	GCGCCAGTAT	GACCAGATCC	TCATCGAAAT	CTGCGGCCAC	AAGGCTATCG	1680
GTACCGTGCT	GGTGGGCCCC	ACACCCGTCA	ACATCATCGG	ACGCAACCTG	TTGACGCAGA	1740
TCGGTTGCAC	GCTGAACTTC	CCCATTAGCC	CTATCGAGAC	GGTACCGGTG	AAGCTGAAGC	1800
CCGGGATGGA	CGGCCCGAAG	GTCAAGCAAT	GGCCATTGAC	AGAGGAGAAG	ATCAAGGCAC	1860
TGGTGGAGAT	TTGCACAGAG	ATGGAAAAGG	AAGGGAAAAT	CTCCAAGATT	GGGCCTGAGA	1920
ACCCGTACAA	CACGCCGGTG	TTGCAATCA	AGAAGAAGGA	CTCGACGAAA	TGGCGCAAGC	1980
TGGTGGACTT	CCGCGAGCTG	AACAAGCGCA	CGCAAGACTT	CTGGGAGGTT	CAGCTGGGCA	2040
TCCCGCACCC	CGCAGGGCTG	AAGAAGAAGA	AATCCGTGAC	CGTACTGGAT	GTGGGTGATG	2100
CCTACTTCTC	CGTTCCCTTG	GACGAAGACT	TCAGGAAGTA	CACTGCCTTC	ACAATCCCTT	2160
CGATCAACAA	CGAGACACCG	GGGATTTCAT	ATCAGTACAA	CGTGCTGCCC	CAGGGCTGGA	2220
AAGGCTCTCC	CGCAATCTTC	CAGAGTAGCA	TGACCAAAAT	CCTGGAGCCT	TTCCGCAAA	2280
AGAACCCCGA	CATCGTCATC	TATCAGTACA	TGGATGACTT	GTACGTGGGC	TCTGATCTAG	2340
AGATAGGGCA	GCACCGCACC	AAGATCGAGG	AGCTGCGCCA	GCACCTGTTG	AGGTGGGGAC	2400
TGACCACACC	CGACAAGAAG	CACCAGAAGG	AGCCTCCCTT	CCTCTGGATG	GGTTACGAGC	2460
TGCACCCTGA	CAAATGGACC	GTGCAGCCTA	TCGTGCTGCC	AGAGAAAGAC	AGCTGGACTG	2520
TCAACGACAT	ACAGAAGCTG	GTGGGGAAGT	TGAAGTGGGC	CAGTCAGATT	TACCCAGGGA	2580
TTAAGGTGAG	GCAGCTGTGC	AAACTCCTCC	CGCGAACCAG	GGCACTCACA	GAGGTGATCC	2640
CCCTAACCGA	GGAGGCCGAG	CTCGAACTGG	CAGAAAACCG	AGAGATCCTA	AAGAGCCCG	2700
TGCACGGCGT	GTACTATGAC	CCCTCCAAGG	ACCTGATCGC	CGAGATCCAG	AAGCAGGGGC	2760
AAGGCCAGTG	GACCTATCAG	ATTTACCAGG	AGCCCTTCAA	GAACCTGAAG	ACCGGCAAGT	2820
ACGCCCGGAT	GAGGGGTGCC	CACACTAACG	ACGTCAAGCA	GCTGACCGAG	GCCGTGCAGA	2880

AGATCACCAC	CGAAAGCATC	GTGATCTGGG	GAAAGACTCC	TAAGTTCAAG	CTGCCCATCC	2940
AGAAGGAAAC	CTGGGAAACC	TGGTGGACAG	AGTATTGGCA	GGCCACCTGG	ATTTCCTGAGT	3000
GGGAGTTCGT	CAACACCCCT	CCCCTGGTGA	AGCTGTGGTA	CCAGCTGGAG	AAGGAGCCCA	3060
TAGTGGGCGC	CGAAACCTTC	TACGTGGATG	GGGCCGCTAA	CAGGGAGACT	AAGCTGGGCA	3120
AAGCCGGATA	CGTCACTAAC	CGGGGCAGAC	AGAAGGTTGT	CACCCTCACT	GACACCACCA	3180
ACCAGAAGAC	TGAGCTGCAG	GCCATTTACC	TCGCTTTGCA	GGACTCGGGC	CTGGAGGTGA	3240
ACATCGTGAC	AGACTCTCAG	TATGCCCTGG	GCATCATTCA	AGCCCAGCCA	GACCAGAGTG	3300
AGTCCGAGCT	GGTCAATCAG	ATCATCGAGC	AGCTGATCAA	GAAGGAAAAAG	GTCTATCTGG	3360
CCTGGGTACC	CGCCCAACAA	GGCATTGGCG	GCAATGAGCA	GGTCGACAAG	CTGGTCTCGG	3420
CTGGCATCAG	GAAGGTGCTA	TTCCTGGATG	GCATCGACAA	GGCCCAGGAC	GAGCACGAGA	3480
AATACCACAG	CAACTGGCGG	GCCATGGCTA	GCGACTTCAA	CCTGCCCCCT	GTGGTGGCCA	3540
AAGAGATCGT	GGCCAGCTGT	GACAAGTGTC	AGCTCAAGGG	CGAAGCCATG	CATGGCCAGG	3600
TGGACTGTAG	CCCCGGCATC	TGGCAACTCG	ATTGCACCCA	TCTGGAGGGC	AAGGTTATCC	3660
TGGTAGCCGT	CCATGTGGCC	AGTGGCTACA	TGAGGCCGA	GGTCATTCCC	GCCGAAACAG	3720
GGCAGGAGAC	AGCCTACTTC	CTCCTGAAGC	TGGCAGGCCG	GTGGCCAGTG	AAGACCATCC	3780
ATACTGACAA	TGGCAGCAAT	TTCACCAGTG	CTACGGTTAA	GGCCGCCTGC	TGGTGGGCGG	3840
GAATCAAGCA	GGAGTTCGGG	ATCCCCCTACA	ATCCCCAGAG	TCAGGGCGTC	GTCTGAGTCTA	3900
TGAATAAGGA	GTTAAAGAAG	ATTATCGGCC	AGGTCAGAGA	TCAGGCTGAG	CATCTCAAGA	3960
CCGCGGTCCA	AATGGCGGTA	TTCATCCACA	ATTTCAGCG	GAAGGGGGGG	ATTGGGGGGT	4020
ACAGTGCGGG	GGAGCGGATC	GTGGACATCA	TCGCGACCGA	CATCCAGACT	AAGGAGCTGC	4080
AAAAGCAGAT	TACCAAGATT	CAGAATTTCC	GGGTCTACTA	CAGGGACAGC	AGAAATCCCC	4140
TCTGGAAAGG	CCCAGCGAAG	CTCCTCTGGA	AGGGTGAGGG	GGCAGTAGTG	ATCCAGGATA	4200
ATAGCGACAT	CAAGGTGGTG	CCCAGAAGAA	AGGCGAAGAT	CATTAGGGAT	TATGGCAAAC	4260
AGATGGCGGG	TGATGATTGC	GTGGCGAGCA	GACAGGATGA	GGATTAG		4307

SEQ. ID. NO. 3 - Envelope Gene from HIV-1 MN (Genbank accession no. M17449)

ATGAGAGTGA	AGGGGATCAG	GAGGAATTAT	CAGCACTGGT	GGGGATGGGG	CACGATGCTC	60
CTTGGGTTAT	TAATGATCTG	TAGTGCTACA	GAAAAATTGT	GGGTCACAGT	CTATTATGGG	120
GTACCTGTGT	GGAAAGAAGC	AACCACCACT	CTATTTTGTG	CATCAGATGC	TAAAGCATAT	180
GATACAGAGG	TACATAATGT	TGGGGCCACA	CAAGCCTGTG	TACCCACAGA	CCCCAACCCA	240
CAAGAAGTAG	AATTGGTAAA	TGTGACAGAA	AATTTTAACA	TGTGGAAAAA	TAACATGGTA	300
GAACAGATGC	ATGAGGATAT	AATCAGTTTA	TGGGATCAAA	GCCTAAAGCC	ATGTGTAAAA	360
TTAACCCAC	TCTGTGTTAC	TTTAAATTGC	ACTGATTTGA	GGAATACTAC	TAATACCAAT	420
AATAGTACTG	CTAATAACAA	TAGTAATAGC	GAGGGAACAA	TAAAGGGAGG	AGAAATGAAA	480
AACTGCTCTT	TCAATATCAC	CACAAGCATA	AGAGATAAGA	TGCAGAAAAG	ATATGCACTT	540
CTTTATAAAC	TTGATATAGT	ATCAATAGAT	AATGATAGTA	CCAGCTATAG	GTTGATAAGT	600
TGTAATACCT	CAGTCATTAC	ACAAGCTTGT	CCAAAGATAT	CCTTTGAGCC	AATTGCCATA	660
CACATATTGT	CCCCGGCTGG	TTTTGCGATT	CTAAAATGTA	ACGATAAAAA	GTTCACTGGA	720
AAAGGATCAT	GTAAAAATGT	CAGCACAGTA	CAATGTACAC	ATGGAATTAG	GCCAGTAGTA	780
TCAACTCAAC	TGCTGTTAAA	TGGCAGTCTA	GCAGAAGAAG	AGGTAGTAAT	TAGATCTGAG	840
AATTTCACTG	ATAATGCTAA	AACCATCATA	GTACATCTGA	ATGAATCTGT	ACAAATTAAT	900
TGTACAAGAC	CCAACACAA	TAAAAGAAAA	AGGATACATA	TAGGACCAGG	GAGAGCATT	960
TATACAACAA	AAAATATAAT	AGGAACATA	AGACAAGCAC	ATTGTAACAT	TAGTAGAGCA	1020
AAATGGAATG	ACACTTTAAG	ACAGATAGTT	AGCAAATTAA	AAGAACAATT	TAAGAATAAA	1080
ACAATAGTCT	TTAATCAATC	CTCAGGAGGG	GACCCAGAAA	TTGTAATGCA	CAGTTTTAAT	1140
TGTGGAGGGG	AATTTTCTA	CTGTAATACA	TCACCACTGT	TTAATAGTAC	TTGGAATGGT	1200
AATAATACTT	GGAATAATAC	TACAGGGTCA	AATAACAATA	TCACACTTCA	ATGCAAAATA	1260
AAACAAATTA	TAAACATGTG	GCAGGAAGTA	GGAAAAGCAA	TGTATGCCCC	TCCCATTGAA	1320
GGACAAATTA	GATGTTTCATC	AAATATTACA	GGGCTACTAT	TAACAAGAGA	TGGTGGTAAG	1380
GACACGGACA	CGAACGACAC	CGAGATCTTC	AGACCTGGAG	GAGGAGATAT	GAGGGACAAT	1440
TGGAGAAGTG	AATTATATAA	ATATAAAGTA	GTAACAATTG	AACCATTAGG	AGTAGCACCC	1500
ACCAAGGCAA	AGAGAAGAGT	GGTGCAGAGA	GAAAAAAGAG	CAGCGATAGG	AGCTCTGTTC	1560
CTTGGGTTCT	TAGGAAGCAG	AGGAAGCACT	ATGGGCGCAG	CGTCAGTGAC	GCTGACGGTA	1620
CAGGCCAGAC	TATTATTGTC	TGGTATAGTG	CAACAGCAGA	ACAATTTGCT	GAGGGCCATT	1680
GAGGCGCAAC	AGCATATGTT	GCAACTCACA	GTCTGGGGCA	TCAAGCAGCT	CCAGGCAAGA	1740
GTCTGGCTG	TGGAAAGATA	CCTAAAGGAT	CAACAGCTCC	TGGGGTTTTG	GGGTTGCTCT	1800
GGAAAACCTCA	TTTGCAACCAC	TACTGTGCCT	TGGAAATGCTA	GTTGGAGTAA	TAAATCTCTG	1860
GATGATATTT	GGAATAACAT	GACCTGGATG	CAGTGGGAAA	GAGAAATTGA	CAATTACACA	1920
AGCTTAATAT	ACTCATTACT	AGAAAAATCG	CAACCCCAAC	AAGAAAAAGAA	TGAACAAGAA	1980
TTATTGGAAT	TGGATAAATG	GGCAAGTTTG	TGGAATTGGT	TTGACATAAC	AAATTGGCTG	2040
TGGTATATAA	AAATATTCAT	AATGATAGTA	GGAGGCTTGG	TAGGTTTAAAG	AATAGTTTTT	2100
GCTGTACTTT	CTATAGTGAA	TAGAGTTAGG	CAGGGATACT	CACCATTGTC	GTTGCAGACC	2160
CGCCCCCAG	TTCCGAGGGG	ACCCGACAGG	CCCGAAGGAA	TGGAAGAAGA	AGGTGGAGAG	2220

AGAGACAGAG	ACACATCCGG	TCGATTAGTG	CATGGATTCT	TAGCAATTAT	CTGGGTGCGAC	2280
CTGCGGAGCC	TGTTCTCTTT	CAGCTACCAC	CACAGAGACT	TACTCTTGAT	TGCAGCGAGG	2340
ATTGTGGAAC	TTCTGGGACG	CAGGGGGTGG	GAAGTCCTCA	AATATTGGTG	GAATCTCCTA	2400
CAGTATTGGA	GTCAGGAAC	AAAGAGTAGT	GCTGTTAGCT	TGCTTAATGC	CACAGCTATA	2460
GCAGTAGCTG	AGGGGACAGA	TAGGGTTATA	GAAGTACTGC	AAAGAGCTGG	TAGAGCTATT	2520
CTCCACATAC	CTACAAGAA	AAGACAGGGC	TTGGAAAGGG	CTTTGCTATA	A	2571

SEQ. I.D. NO. 4 - SYNgp-160mn - codon optimised env sequence

ATGAGGGTGA	AGGGGATCCG	CCGCAACTAC	CAGCACTGGT	GGGGCTGGGG	CACGATGCTC	60
CTGGGGCTGC	TGATGATCTG	CAGCGCCACC	GAGAAGCTGT	GGGTGACCGT	GTACTACGGC	120
GTGCCCCTGT	GGAAGGAGGC	CACCACCACC	CTGTTCTGCG	CCAGCGACGC	CAAGGCGTAC	180
GACACCGAGG	TGCACAACGT	GTGGGGCCACC	CAGGCGTGCG	TGCCCCACCGA	CCCCAACCCC	240
CAGGAGGTGG	AGCTCGTGAA	CGTGACCAG	AACCTTCAACA	TGTGGAAGAA	CAACATGGTG	300
GAGCAGATGC	ATGAGGACAT	CATCAGCCTG	TGGGACCAGA	GCCTGAAGCC	CTGCGTGAAG	360
CTGACCCCCC	TGTGCGTGAC	CCTGAACCTG	ACCGACCTGA	GGAACACCAC	CAACACCAAC	420
AACAGCACCG	CCAACAACAA	CAGCAACAGC	GAGGGCACCA	TCAAGGGCGG	CGAGATGAAG	480
AACCTGCAGCT	TCAACATCAC	CACCAGCATC	CGCGACAAGA	TGCAGAAGGA	GTACGCCCTG	540
CTGTACAAGC	TGGATATCGT	GAGCATCGAC	AACGACAGCA	CCAGCTACCG	CCTGATCTCC	600
TGCAACACCA	GCGTGATCAC	CCAGGCTGCG	CCCAAGATCA	GCTTCGAGCC	CATCCCCATC	660
CACCTACTGCG	CCCCCGCCGG	CTTCGCCATC	CTGAAGTGCA	ACGACAAGAA	GTTTCAGCGG	720
AAGGGCAGCT	GCAAGAACGT	GAGCACCGTG	CAGTGACCCC	ACGGCATCCG	GCCGGTGGTG	780
AGCACCCAGC	TCCTGCTGAA	CGGCAGCCTG	GCCGAGGAGG	AGGTGGTGAT	CCGCAGCGAG	840
AACCTTACCG	ACAACGCCAA	GACCATCATC	GTGCACCTGA	ATGAGAGCGT	GCAGATCAAC	900
TGCACGCGTC	CCAACCTACAA	CAAGCGCAAG	CGCATCCACA	TCGGCCCCCG	GCGCGCCTTC	960
TACACCACCA	AGAACATCAT	CGGCACCATC	CGCCAGGCCC	ACTGCAACAT	CTCTAGAGCC	1020
AAGTGGAAAC	ACACCTTGCG	CCAGATCGTG	AGCAAGCTGA	AGGAGCAGTT	CAAGAACAAG	1080
ACCATCGTGT	TCAACCAGAG	CAGCGGCGGC	GACCCCGAGA	TCGTGATGCA	CAGCTTCAAC	1140
TGCGGCGGCG	AATTCTTCTA	CTGCAACACC	AGCCCCCTGT	TCAACAGCAC	CTGGAACGGC	1200
AACAACACCT	GGAACAACAC	CACCGGCAGC	AACAACAATA	TTACCCTCCA	GTGCAAGATC	1260
AAGCAGATCA	TCAACATGTG	GCAGGAGGTG	GGCAAGGCCA	TGTACGCCCC	CCCCATCGAG	1320
GGCCAGATCC	GGTGACAGC	CAACATCACC	GGTCTGCTGC	TGACCCGCGA	CGGCGGCAAG	1380
GACACCGACA	CCAAACGAC	CGAAATCTTC	CGCCCCGGCG	GCGGCGACAT	GCGGACAAAC	1440
TGGAGATCTG	AGCTGTACAA	GTACAAGGTG	GTGACGATCG	AGCCCCCTGG	CGTGGCCCCC	1500
ACCAAGGCCA	AGCGCCGCGT	GGTGACGCGC	GAGAAGCGGG	CCGCCATCGG	CGCCCTGTTC	1560
CTGGGCTTCC	TGGGGGCGGC	GGGCAGCACC	ATGGGGGCGG	CCAGCGTGAC	CCTGACCGTG	1620
CAGGCCCCCG	TGCTCCTGAG	CGGCATCGTG	CAGCAGCAGA	ACAACCTCCT	CCGCGCCATC	1680
GAGGCCCCAG	AGCATATGCT	CCAGCTCACC	GTGTGGGGCA	TCAAGCAGCT	CCAGGCCCCG	1740
GTGCTGGCCG	TGGAGCGCTA	CCTGAAGGAC	CAGCAGCTCC	TGGGCTTCTG	GGGCTGTCTC	1800
GGCAAGCTGA	TCTGCACCAC	CACGGTACCC	TGGAACGCCT	CCTGGAGCAA	CAAGAGCCTG	1860
GACGACATCT	GGAACAACAT	GACCTGGATG	CAGTGGGAGC	GCGAGATCGA	TAACCTACAC	1920
AGCCTGATCT	ACAGCCTGCT	GGAGAAGAGC	CAGACCCAGC	AGGAGAAGAA	CGAGCAGGAG	1980
CTGCTGGAGC	TGGACAAGTG	GGCGAGCCTG	TGGAACCTGG	TCGACATCAC	CAACTGGCTG	2040
TGGTACATCA	AAATCTTCAT	CATGATTGTG	GGCGGCTTGG	TGGGCTTCCG	CATCGTGTTC	2100
GCCGTGCTGA	GCATCGTGAA	CCGCGTGCGC	CAGGGCTACA	GCCCCCTGAG	CCTCCAGACC	2160
CGGCCCCCGG	TGCCGCGCGG	GCCCGACCGC	CCCGAGGGCA	TCGAGGAGGA	GGGCGGCGAG	2220
CGCGACCGCG	ACACCAGCGG	CAGGCTCGTG	CACGGCTTCC	TGGCGATCAT	CTGGGTGCGAC	2280
CTCCGCAGCC	TGTTCTCTGT	CAGCTACCAC	CACCGCGACC	TGCTGCTGAT	CGCCGCCCCG	2340
ATCGTGGAA	TCCTAGGCCG	CCGCGGCTGG	GAGGTGCTGA	AGTACTGGTG	GAACCTCCTC	2400
CAGTATTGGA	GCCAGGAGCT	GAAGTCCAGC	CCCGTGAGCC	TGCTGAACGC	CACCGCCATC	2460
GCCGTGGCCG	AGGGCACCGA	CCGCGTGATC	GAGGTGCTCC	AGAGGGCCGG	GAGGGCGATC	2520
CTGCACATCC	CCACCCGCAT	CCGCCAGGGG	CTCGAGAGGG	CGCTGCTGTA	A	2571

SEQ. I.D. NO. 11 - Complete Sequence of pH4DOZENEGS

CTGACGCGCC	CTGTAGCGGC	GCATTAAGCG	CGGCGGGTGT	GGTGGTTACG	CGCAGCGTGA	60
CCGCTACACT	TGCCAGCGCC	CTAGCGCCCC	CTCCTTTCGC	TTTCTTCCCT	TCCTTTCTCG	120
CCACGTTTCG	CGGCTTTCCC	CGTCAAGCTC	TAAATCGGGG	GCTCCCTTTA	GGGTTCCGAT	180
TGAGTGCTTT	ACGGCACCTC	GACCCCAAAA	AACTTGATTA	GGGTGATGGT	TCACGTAGTG	240
GGCCATCGCC	CTGATAGACG	GTTTTTCGCC	CTTTGACGTT	GGAGTCCACG	TTCTTTAATA	300
GTGGACTCTT	GTTCCAAACT	GGAACAACAC	TCAACCCAT	CTCGGTCTAT	CTTTTTGATT	360
TATAAGGGAT	TTTGCCGATT	TCGGCCTATT	GGTTAAAAAA	TGAGCTGATT	TAACAAAAAT	420
TTAACGCGAA	TTTTAACAAA	ATATTAACGC	TTACAATTTT	CATTCGCCAT	TCAGGCTGCG	480

CAACTGTTGG GAAGGGCGAT CGGTGCGGGC CTCTTCGCTA TTACGCCAGC TGGCGAAAGG 540
 GGGATGTGCT GCAAGGCGAT TAAGTTGGGT AACGCCAGGG TTTTCCCAGT CACGACGTTG 600
 TAAAACGACG GCCAGTGAGC GCGCGTAATA CGACTCACTA TAGGGCGAAT TGGAGCTCCA 660
 CCGCGGTGGC GGCCGCTCTA GAGTCCGTTA CATAACTTAC GGTAAATGGC CCGCCTGGCT 720
 GACCGCCCAA CGACCCCGC CCATTGACGT CAATAATGAC GTATGTTCCC ATAGTAACGC 780
 CAATAGGGAC TTTCCATTGA CGTCAATGGG TGGAGTATTT ACGGTAAACT GCCCACTTGG 840
 CAGTACATCA AGTGTATCAT ATGCCAAGTA CGCCCCCTAT TGACGTCAAT GACGGTAAAT 900
 GGCCCGCCTG GCATTATGCC CAGTACATGA CCTTATGGGA CTTTCTACT TGGCAGTACA 960
 TCTACGTATT AGTCATCGCT ATTACCATGG TGATGCGGTT TTGGCAGTAC ATCAATGGGC 1020
 GTGGATAGCG GTTTGACTCA CGGGGATTTT CAAGTCTCCA CCCCATTGAC GTCAATGGGA 1080
 GTTTGTTTTG GCACCAAAAT CAACGGGACT TTCCAAAATG TCGTAACAAC TCCGCCCCAT 1140
 TGACGCAAAAT GGGCGGTAGG CGTGTACGGT GGGAGGTCTA TATAAGCAGA GCTCGTTTTAG 1200
 TGAACCGGTC TCTCTGGTTA GACCAGATCT GAGCTTGGGA GCTCTCTGGC TAACTCAGGC 1260
 ACCCACTGCT TAAGCCTCAA TAAAGCTTGC CTTGAGTGCT TCAAGTAGTG TGTGCCCGTC 1320
 TGTTGTGTGA CTCTGGTAAC TAGAGATCCC TCAGACCCTT TTAGTCAGTG TGGAAAATCT 1380
 CTAGCAGTGG CGCCCGAACA GGGACTTGAA AGCGAAAGGG AAACCAGAGG AGCTCTCTCG 1440
 ACGCAGGACT CGGCTTGCTG AAGCGCGCAC GGCAAGAGGC GAGGGGCGGC GACTGGTGAG 1500
 TACGCCAAAA ATTTTGACTA GCGGAGGCTA GAAGGAGAGA GATGGGTGCG AGAGCGTCAG 1560
 TATTAAGCGG GGGAGAATTA GATCGCGATG GGAAAAAATT CGGTTAAGGC AGCGGGGAAA 1620
 GAAAAAATAT AAATTAAAAC ATATAGTATG GGCAAGCAGG GAGCTAGAAC GATTCGCAGT 1680
 TAATCCTGGC CTGTTAGAAA CATCAGAAGG CTGTAGACAA ATACTGGGAC AGCTACAACC 1740
 ATCCCTTCAG ACAGGATCAG AAGAACTTAG ATCATTATAT AATACAGTAG CAACCCTCTA 1800
 TTGTGTGCAT CAAAGGTTGA GATAAAAGAC ACCAAGGAAG CTTTAGACAA GATAGAGGGA 1860
 GAGCAAAACA AAAGTAAGAA AAAAGCACAG CAAGCAGCAG CTGACACAGG ACACAGCAAT 1920
 CAGGTCAGCC AAAATTATCC TATAGTGCAG AACATCCAGG GGCAAATGGT ACATCAGGCC 1980
 ATATCACCTA GAACTTTAAA TGCATGGGTA AAAGTAGTAG AAGAGAAGGC TTTCAGCCCA 2040
 GAAGTGATAC CCATGTTTTT AGCATTATCA GAAGGAGCCA CCCACAAGA TTTAAACACC 2100
 ATGCTAAACA CAGTGGGGGG ACATCAAGCA GCCATGCAAA TGTTAAAAGA GACCATCAAT 2160
 GAGGAAGCTG CAGGAATTCG CCTAAAACCT CTTGTACCAA TTGCTATTGT AAAAAGTGTT 2220
 GCTTTTCATT CCAAGTTTGT TTCATAACAA AAGCCTTAGG CATCTCCTAT GGCAGGAAGA 2280
 AGCGGAGACA GCGACGAAGA GCTCATCAGA ACAGTCAGAC TCATCAAGCT TCTCTATCAA 2340
 AGCAGTAAGT AGTACATGTA ACGCAACCTA TACCAATAGT AGCAATAGTA GCATTAGTAG 2400
 TAGCAATAAT AATAGCAATA GTTGTGTGGT CCATAGTAAT CATAGAATAT AGGAAAATAT 2460
 TAAGACAAAG AAAAATAGAC AGGTTAATTG ATAGACTAAT AGAAAGAGCA GAAGACAGTG 2520
 GCAATGAGAG TGAAGGAGAA ATATCAGCAC TTGTGGAGAT GGGGGTGGAG ATGGGGCACC 2580
 ATGCTCCTTG GATGTGTGAT GATCTGTAGT GCTACAGAAA AATTGTGGGT CACAGTCTAT 2640
 TATGGGGTAC CTGTGTGGAA GGAAGCAACC ACCACTCTAT TTTGTGCAT AGATGCTAAA 2700
 GCATAGATCT TCAGACTTGG AGGAGGAGAT ATGAGGGACA ATTGGAGAAG TGAATTATAT 2760
 AAATATAAAG TAGTAAAAAT TGAACCATTA GGAGTAGCAC CCACCAAGGC AAAGAGAAGA 2820
 GTGGTGCAGA GAGAAAAAAG AGCAGTGGGA ATAGGAGCTT TGTTCTTGG GTTCTTGGGA 2880
 GCAGCAGGAA GCACTATGGG CGCAGCGTCA ATGACGCTGA CGGTACAGGC CAGACAATTA 2940
 TTGTCTGGTA TAGTGCAGCA GCAGAACAAT TTGCTGAGG CTATTGAGG GCAACAGCAT 3000
 CTGTTGCAAC TCACAGTCTG GGGCATCAAG CAGCTCCAGG CAAGAATCCT GGCTGTGGAA 3060
 AGATACCTAA AGGATCAACA GCTCCTGGGG ATTTGGGGTT GCTCTGGAAG ACTCATTTGC 3120
 ACCACTGCTG TGCCCTGGAA TGCTAGTTGG AGTAATAAAT CTCTGGAACA GATCTGGAAT 3180
 CACACGACCT GGATGGAGTG GGACAGAGAA ATTAACAATT ACACAAGCTT AATACACTCC 3240
 TTAATTGAAG AATCGCAAAA CCAGCAAGAA AAGAATGAAC AAGAATTATT GGAATTAGAT 3300
 AAATGGGCAA GTTTGTGGAA TTGGTTTAAC ATAACAAATT GGCTGTGGTA TATAAATTA 3360
 TTCATAATGA TAGTAGGAGG CTTGGTAGGT TTAAGAATAG TTTTGTGCT ACTTTCTATA 3420
 GTGAATAGAG TTAGGCAGGG ATATTACCA TTATCGTTT AGACCCACCT CCCAACCCCG 3480
 AGGGGACCCG ACAGGCCCGA AGGAATAGAA GAAGAAGGTG GAGAGAGAGA CAGAGACAGA 3540
 TCCATTGCGT TAGTGAACGG ATCCTTGGCA CTTATCTGGG ACGATCTGCG GAGCCTGTGC 3600
 CTCTTCAGCT ACCACGCTT GAGAGACTTA CTCTTGATTG TAACGAGGAT TGTGGAACTT 3660
 CTGGGACGCA GGGGGTGGGA AGCCCTCAAA TATTGTGGTA ATCTCTACA GTATTGGAGT 3720
 CAGGAATAA AGAATAGTGC TGTTAGCTTG CTCAATGCCA CAGCCATAGC AGTAGCTGAG 3780
 GGGACAGATA GGGTTATAGA AGTAGTACAA GGAGCTTGTA GAGCTATTTC CCACATACCT 3840
 AGAAGAATAA GACAGGGCTT GGAAAGGATT TTGCTATAAG ATGGGTGGCA AGTGGTCAAA 3900
 AAGTAGTGTG ATTGGATGGC CTACTGTAA GGAAGAATG AGACGAGCTG AGCCAGCAGC 3960
 AGATAGGGTG GAGAGCAGAT CTCGACGCTG CAGGAGTGGG GAGGCACGAT GGCCGCTTTG 4020
 GTCGAGGCGG ATCCGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT AAATCAATAT 4080
 TGGCTATTGG CCATTGCATA CGTTGTATCC ATATCATAAT ATGTACATTT ATATTGGCTC 4140
 ATGTCCAACA TTACCGCCAT GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT 4200
 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTCCGCG GTTACATAAC TTACGGTAAA 4260
 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT 4320
 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA 4380

AACTGCCCCAC	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	AGTACGCCCC	CTATTGACGT	4440
CAATGACGGT	AAATGGCCCC	CCTGGCATT	TGCCCCAGTAC	ATGACCTTAT	GGGACTTTCC	4500
TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATGC	GGTTTTGGCA	4560
GTACATCAAT	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC	TCCACCCCAT	4620
TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA	AATGTCGTAA	4680
CAACTCCGCC	CCATTGACGC	AAATGGGCGG	TAGGCATGTA	CGGTGGGAGG	TCTATATAAG	4740
CAGAGCTCGT	TTAGTGAACC	GTCAGATCGC	CTGGAGACGC	CATCCACGCT	GTTTTGACCT	4800
CCATAGAAGA	CACCGGGACC	GATCCAGCCT	CCGCGGCCCC	AAGCTTCAGC	TGCTCGAGCC	4860
CGGGGATGAC	GTCATCGACT	TCGAAGGTTT	GAATCCTTCT	ACTGCCACCA	TTTTTTCTCT	4920
ACGTCATCGA	CTTCGAAGGT	TCGAATCCTT	CCCTGTCCAC	CAGTCGAGTA	TTACGTCATC	4980
GACTTCGAAG	GTTCGAATCC	TTCTAGATT	ACCAATTTT	AGGAACGTCA	TCGACTTCGA	5040
AGGTTTCGAAT	CCTTCCAGTT	CCACCACTCG	AGGCCAACGT	CATCGACTTC	GAAGGTTTCA	5100
ATCCTTCTCT	TCCCACCATT	TTTTTTCCAC	GTCATCGACT	TCGAAGGTTT	GAATCCTTCG	5160
GGGCCCCACCA	GTCGAGGGCT	ACGTCATCGA	CTTCGAAGGT	TCGAATCCTT	CTTGCTTCAC	5220
CATTTTTTCT	GAACGTCATC	GACTTCGAAG	GTTTCGAATCC	TTCTGCTGTC	ACCAGTCGAG	5280
TATAACGTCA	TCGACTTCGA	AGGTTTCAAT	CCTTCACCGG	TCACCATTTT	TTTATAACGT	5340
CATCGACTTC	GAAGGTTTCA	ATCCTTCTTC	TTACACCACT	CGAGGTACAC	GTCATCGACT	5400
TCGAAGGTTT	GAATCCTTCG	TAGTTTACCA	TTTTTTGTGC	ACGTCATCGA	CTTCGAAGGT	5460
TCGAATCCTT	CTAGGCCCCAC	CAGTCGACGC	ATGCCTGCAG	GTCGAGGTCG	ATACCGTCGA	5520
GACCTAGAAA	AACATGGAGC	AATCACAAAGT	AGCAATACAG	CAGCTACCAA	TGCTGATTGT	5580
GCCTGGCTAG	AAGCACAAGA	GGAGGAGGAG	GTGGGTTTTT	CAGTCACACC	TCAGGTACCT	5640
TTAAGACCAA	TGACTTACAA	GGCAGCTGTA	GATCTTAGCC	ACTTTTTTAA	AGAAAAGGGG	5700
GGACTGGAAG	GGCTAATTCA	CTCCCAACGA	AGACAAGATA	TCCTTGATCT	GTGGATCTAC	5760
CACACACAAG	GCTACTTCCC	TGATTGGCAG	AACACACAC	CAGGGCCAGG	TAGACATAT	5820
CCACTGACCT	TTGGATGGTG	CTACAAGCTA	GTACCAGTTG	AGCAAGAGAA	GGTAGAAGAA	5880
GCCAATGAAG	GAGAGAACAC	CCGCTTGTTA	CACCCTGTGA	GCCTGCATGG	GATGGATGAC	5940
CCGGAGAGAG	AAGTATTAGA	GTGGAGGTTT	GACAGCCGCC	TAGCATTTCA	TCACATGGCC	6000
CGAGAGCTGC	ATCCGGAGTA	CTTCAAGAAC	TGCTGACATC	GAGCTTGCTA	CAAGGGACTT	6060
TCCGCTGGGG	ACTTTCAGG	GAGGCGTGCC	CTGGGCGGGA	CTGGGGAGTG	GCGAGCCCTC	6120
AGATGCTGCA	TATAAGCAGC	TGCTTTTTGC	CTGTACTGGG	TCTCTCTGGT	TAGACCAGAT	6180
CTGAGCCTGG	GAGCTCTCTG	GCTAACTAGG	GAACCCACTG	CTTAAGCCTC	AATAAAGCTT	6240
GCCTTGAGTG	CTTCAAGTAG	TGTGTGCCCC	TCTGTTGTGT	GACTCTGGTA	ACTAGAGATC	6300
CCTCAGACCC	TTTTAGTCAG	TGTGGAAAAT	CTCTAGCAGT	CGAGGGGGGG	CCCGGTACCC	6360
AGCTTTTGT	CCCTTTAGT	AGGGTTAAT	GCGCGCTTGG	CGTAATCATG	GTCATAGCTG	6420
TTTCTGTGT	AAAATTGTTA	TCCGCTCACA	ATTCCACACA	ACATACGAGC	CGGAAGCATA	6480
AAGTGTAAG	CCTGGGGTGC	CTAATGAGTG	AGCTAACTCA	CATTAATTGC	GTTGCGCTCA	6540
CTGCCCCGCT	TCCAGTCGGG	AAACCTGTGC	TGCCAGCTGC	ATTAATGAAT	CGGCCAACGC	6600
GCGGGGAGAG	GCGGTTTGCG	TATTGGGCGC	TCTTCCGCTT	CCTCGCTCAC	TGACTCGCTG	6660
CGCTCGGTGC	TTCGGCTGCG	GCGAGCGGTA	TCAGCTCACT	CAAAGGCGGT	AATACGGTTA	6720
TCCACAGAAT	CAGGGGATAA	CGCAGGAAAG	AACATGTGAG	CAAAAGGCCA	GCAAAAGGCC	6780
AGGAACCGTA	AAAAGGCCGC	GTTGCTGGCG	TTTTTCCATA	GGCTCCGCC	GGCTAGCGAG	6840
CATCACAAAA	ATCGACGCTC	AAGTCAGAGG	TGGCGAAACC	CGACAGGACT	ATAAAGATAC	6900
CAGGCGTTTC	CCCCTGGAAG	CTCCCTCGTG	CGCTCTCCTG	TTCCGACCCT	GCCGCTTACC	6960
GGATACCTGT	CCGCCTTTCT	CCCTTCGGGA	AGCGTGGCGC	TTTCTCATAG	CTCACGCTGT	7020
AGGTATCTCA	GTTCCGGTGT	GGTCGTTGCG	TCCAAGCTGG	GCTGTGTGCA	CGAACCCCCC	7080
GTTACGCCCC	ACCGCTGCGC	CTTATCCGGT	AACATCGTC	TTGAGTCCAA	CCCGGTAAGA	7140
CACGACTTAT	CGCCACTGGC	AGCAGCCACT	GGTAACAGGA	TTAGCAGAGC	GAGGTATGTA	7200
GGCGGTGCTA	CAGAGTTCTT	GAAGTGGTGG	CCTAACTACG	GCTACACTAG	AAGGACAGTA	7260
TTTGGTATCT	GCGCTCTGCT	GAAGCCAGTT	ACCTTCGGAA	AAAGAGTTGG	TAGCTCTTGA	7320
TCCGGCAAAAC	AAACCACCGC	TGGTAGCGGT	GGTTTTTTTT	TTTGCAAGCA	GCAGATTACG	7380
CGCAGAAAAA	AAGGATCTCA	AGAAGATCCT	TTGATCTTTT	CTACGGGGTC	TGACGCTCAG	7440
TGGAACGAAA	ACTCACGTTA	AGGGATTTTG	GTCATGAGAT	TATCAAAAAG	GATCTTCACC	7500
TAGATCCTTT	TAAATTAAAA	ATGAAGTTTT	AAATCAATCT	AAAGTATATA	TGAGTAAACT	7560
TGGTCTGACA	GTTACCAATG	CTTAATCAGT	GAGGCACCTA	TCTCAGCGAT	CTGTCTATTT	7620
CGTTTATCCA	TAGTTGCCTG	ACTCCCCGTC	GTGTAGATAA	CTACGATACG	GGAGGGCTTA	7680
CCATCTGGCC	CCAGTGCTGC	AATGATACCG	CGAGACCCAC	GCTCACCGGC	TCCAGATTTA	7740
TCAGCAATAA	ACCAGCCAGC	CGGAAGGGCC	GAGCGCAGAA	GTGGTCTGTC	AACTTTATCC	7800
GCCTCCATCC	AGTCTATTAA	TTGTTGCCGG	GAAGTATAG	TAAGTAGTTT	CCGAGTAAAT	7860
AGTTTGCGCA	ACGTTGTTGC	CATTGCTACA	GGCATCGTGG	TGTCACGCTC	GTCGTTTGGT	7920
ATGGCTTCAT	TCAGCTCCGG	TTCCCAACGA	TCAAGGCGAG	TTACATGATC	CCCCATGTTG	7980
TGCAAAAAAG	CGGTTAGCTC	CTTCGGTCCT	CCGATCGTTG	TCAGAAGTAA	GTTGGCCGCA	8040
GTGTTATCAC	TCATGGTTAT	GGCAGCACTG	CATAATTCTC	TTACTGTCTA	GCCATCCGTA	8100
AGATGCTTTT	CTGTGACTGG	TGAGTACTCA	ACCAAGTCA	TCTGAGAATA	GTGTATGCGG	8160
CGACCGAGTT	GCTCTTGCCC	GGCGTCAATA	CCGGGATAATA	CCGCGCCACA	TAGCAGAACT	8220
TTAAAAGTGC	TCATCATTGG	AAAACGTTCT	TCGGGGCGAA	AACTCTCAAG	GATCTTACCG	8280

CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT 8340
 ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA 8400
 ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC TTTTTCATA TTATTGAAGC 8460
 ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA GAAAAATAAA 8520
 CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC 8560

SEQ. I.D. NO. 12 - pSYNGP2 - codon optimised HIV-1 gagpol with leader sequence

1 GGGTCTCTCT GGTTAGACCA GATCTGAGCC TGGGAGCTCT CTGGCTAACT AGGGAACCCA
 61 CTGCTTAAGC CTCAATAAAG CTTGCCTTGA GTGCTTCAAG TAGTGTGTGC CCGTCTGTTG
 121 TGTGACTCTG GTAAGTAGAG ATCCCTCAGA CCCTTTTAGT CAGTGTGGAA AATCTCTAGC
 181 AGTGGCGCCC GAACAGGGAC CTGAAAGCGA AAGGGAAACC AGAGCTCTCT CGACGCAGGA
 241 CTCGGCTTGC TGAAGCGCCC GCACGGCAAG AGGCGAGGGG CGGCGACTGG TGAGTACGCC
 301 AAAAATTTTG ACTAGCGGAG GCTAGAAGGA GAGAGATGGG CGCCCGCGCC AGCGTGCTGT
 361 CGGGCGGCGA GCTGGACCGC TGGGAGAAGA TCCGCCTGCG CCGCGGCGGC AAAAAGAAGT
 421 ACAAGCTGAA GCACATCGTG TGGGCCAGCC GCGAACTGGA GCGCTTCGCC GTGAACCCCG
 481 GGCTCCTGGA GACCAGCGAG GGGTGCCGCC AGATCCTCGG CCAACTGCAG CCCAGCCTGC
 541 AAACCGGCAG CGAGGAGCTG CGCAGCCTGT ACAACACCGT GGCCACGCTG TACTGCGTCC
 601 ACCAGCGCAT CGAAATCAAG GATACGAAAG AGGCCCTGGA TAAATCGAA GAGGAACAGA
 661 ATAAGAGCAA AAAGAAGGCC CAACAGGCCG CCGCGGACAC CGGACACAGC AACCAGGTCA
 721 GCCAGAATA CCCCATCGTG CAGAACATCC AGGGGCAGAT GGTGCACAGC GCCATCTCCC
 781 CCCGCACGCT GAACGCCTGG GTGAAGGTGG TGAAGAGAA GGCTTTTAGC CCGGAGGTGA
 841 TACCCATGTT CTCAGCCCTG TCAGAGGGAG CCACCCCCCA AGATCTGAAC ACCATGCTCA
 901 ACACAGTGGG GGGACACCAG GCCGCCATGC AGATGCTGAA GGAGACCATC AATGAGGAGG
 961 CTGCCGAATG GGATCGTGTG CATCCGGTGC ACGCAGGGCC CATCGCACCG GGCCAGATGC
 1021 GTGAGCCACG GGGCTCAGAC ATCGCCGGAA CGACTAGTAC CCTTCAGGAA CAGATCGGCT
 1081 GGATGACCAA CAACCCACCC ATCCCGGTGG GAGAAATCTA CAAACCTGG ATCATCCTGG
 1141 GCCTGAACAA GATCGTGCGC ATGTATAGCC CTACCAGCAT CCTGGACATC CGCCAAGGCC
 1201 CGAAGGAACC CTTTCGCGAC TACGTGGACC GGTTCACAA AACGCTCCGC GCCGAGCAGG
 1261 CTAGCCAGGA GGTGAAGAAC TGGATGACCG AAACCTGCT GGTCCAGAAC GCGAACCCTGG
 1321 ACTGCAAGAC GATCCTGAAG GCCCTGGGCC CAGCGGCTAC CCTAGAGGAA ATGATGACCG
 1381 CCTGTAGGG AGTGGGCGGA CCCGGCCACA AGGCACGCGT CCTGGCTGAG GCCATGAGCC
 1441 AGGTGACCAA CTCGCTACC ATCATGATGC AGCGCGGCAA CTTTCGGAAC CAACGCAAGA
 1501 TCGTCAAGTG CTTCAACTGT GGCAAGAAG GGCACACAGC CCGCAACTGC AGGGCCCCCTA
 1561 GGAAAAAGGG CTGTTGGAAG TGTGGAAGG AAGGACACCA AATGAAAGAT TGTAAGTGA
 1621 GACAGGCTAA TTTTFTAGGG AAGATCTGGC CTTCCACAA GGGAAAGGCC GGGAAATTTTC
 1681 TTCAGAGCAG ACCAGAGCCA ACAGCCCCAC CAGAAGAGAG CTTCAGGTTT GGGGAAGAGA
 1741 CAACAACCTC CTCACGAAG CAGGAGCCGA TAGACAAGGA ACTGTATCCT TTAGCTTCCC
 1801 TCAGATCACT CTTTGGCAGC GACCCCTCGT CACAATAAAG ATAGGGGGGC AGCTCAAGGA
 1861 GGCTCTCCTG GACACCGGAG CAGACGACAC CGTGCTGGAG GAGATGTCTG TGCCAGGCCG
 1921 CTGGAAGCCG AAGATGATCG GGGGAATCGG CGGTTTCATC AAGGTGCGCC AGTATGACCA
 1981 GATCCTCATC GAAATCTGCG GCCACAAGGC TATCGGTACC GTGCTGGTGG GCGCCACACC
 2041 CGTCAACATC ATCGGACGCA ACCTGTTGAC GCAGATCGGT TGCACGCTGA ACTTCCCCAT
 2101 TAGCTTCTATC GAGACGGTAC CGGTGAAGCT GAAGCCCGG ATGGACGGCC CGAAGGTCAA
 2161 GCAATGGCCA TTGACAGAGG AGAAGATCAA GGCACCTGGT GAGATTGCA CAGAGATGGA
 2221 AAAGGAAGGG AAAATCTCCA AGATTGGGCC TGAGAACCCG TACAACACGC CGGTGTTCTGC
 2281 AATCAAGAAG AAGGACTCGA CGAAATGGCG CAAGCTGGTG GACTTCCGCG AGCTGAACAA
 2341 GCGCACGCAA GACTTCTGGG AGGTTTCACT GGGCATCCCG CACCCCGCAG GGCTGAAGAA
 2401 GAAGAAATCC GTGACCGTAC TGGATGTGGG TGATGCCTAC TTCTCCGTTT CCTGGGACGA
 2461 AGACTTCAGG AAGTACACTG CTTTCACAAT CCCTTCGATC AACAACGAGA CACCGGGGAT
 2521 TCGATATCAG TACAACGTGC TGCCCCAGGG CTGGAAAGGC TCTCCCGCAA TCTTCCAGAG
 2581 TAGCATGACC AAAATCCTGG AGCCTTTCCG CAAACAGAAC CCCGACATCG TCATCTATCA
 2641 GTACATGGAT GACTTGTACG TGGGCTCTGA TCTAGAGATA GGGCAGCACC GCACCAAGAT
 2701 CGAGGAGCTG CGCCAGCACC TGTTGAGGTG GGGACTGACC ACACCCGACA AGAAGCACCA
 2761 GAAGGAGCCT CCCTTCTCTT GAGTGGGTTA CGAGCTGCAC CCTGACAAAT GGACCGTGCA
 2821 GCCTATCGTG CTGCCAGAGA AAGACAGCTG GACTGTCAAC GACATACAGA AGCTGGTGGG
 2881 GAAGTTGAAC TGGGCCAGTC AGATTTACCC AGGGATTAAAG GTGAGGCAGC TGTGCAAACT
 2941 CCTCCGCGGA ACCAAGGCAC TCACAGAGGT GATCCCCCTA ACCGAGGAGG CCGAGCTCGA
 3001 ACTGGCAGAA AACCAGAGGA TCCTAAAGGA GCCCGTGCAC GCGGTGTACT ATGACCCCTC
 3061 CAAGGACCTG ATCGCCGAGA TCCAGAAGCA GGGGCAAGGC CAGTGGACCT ATCAGATTTA
 3121 CCAGGAGCCC TTCAAGAACC TGAAGACCGG CAAAGTACGCC CGGATGAGGG GTGCCACAC
 3181 TAACGACGTC AAGCAGCTGA CCGAGGCCGT GCAGAAAGATC ACCACCGAAA GCATCGTGAT
 3241 CTGGGGAAAG ACTCCTAAGT TCAAGCTGCC CATCCAGAAG GAAACCTGGG AAACCTGGTG
 3301 GACAGAGTAT TGGCAGGCCA CCTGGATTCC TGAGTGGGAG TTCGTCAACA CCCCTCCCTC
 3361 GGTGAAGCTG TGGTACCAGC TGGAGAAGGA GCCCATAGTG GGCGCCGAAA CTTTCTACGT

09936572.121101

```

3421 GGATGGGGCC GCTAACAGGG AGACTAAGCT GGGCAAAGCC GGATACGTCA CTAACCGGGG
3481 CAGACAGAAG GTTGTACACC TCACTGACAC CACCAACCAG AAGACTGAGC TGCAGGCCAT
3541 TTACCTCGCT TTGCAGGACT CGGGCCTGGA GGTGAACATC GTGACAGACT CTCAGTATGC
3601 CCTGGGCATC ATTCAGGCC AGCCAGACCA GAGTGAGTCC GAGCTGGTCA ATCAGATCAT
3661 CGAGCAGCTG ATCAAGAAGG AAAAGGTCTA TCTGGCCTGG GTACCCGCCC ACAAAAGGCAT
3721 TGGCGGCAAT GAGCAGGTCG ACAAGCTGGT CTCGGCTGGC ATCAGGAAGG TGCTATTCCCT
3781 GGATGGCATC GACAAGGCC AGGACGAGCA CGAGAAATAC CACAGCAACT GCGGGGCCAT
3841 GGCTAGCGAC TTCAACCTGC CCCCTGTGGT GGCCAAAGAG ATCGTGGCCA GCTGTGACAA
3901 GTGTCAAGTC AAGGGCGAAG CCATGCATGG CCAGGTGGAC TGTAGCCCCG GCATCTGGCA
3961 ACTCGATTGC ACCCATCTGG AGGGCAAGGT TATCCTGGTA GCCGTCCATG TGGCCAGTGG
4021 CTACATCGAG GCCGAGTCA TTCCCGCCGA AACAGGGCAG GAGACAGCCT ACTTCCTCCT
4081 GAAGCTGGCA GGCCGGTGGC CAGTGAAGAC CATCCATACT GACAATGGCA GCAATTTTAC
4141 CAGTGCTACG GTTAAGGCCG CCTGCTGGTG GGCGGGAATC AAGCAGGAGT TCGGGATCCC
4201 CTACAATCCC CAGAGTCAGG GCGTCGTCGA GTCTATGAAT AAGGAGTTAA AGAAGATTAT
4261 CGGCCAGGTC AGAGATCAGG CTGAGCATCT CAAGACCGCG GTCCAAATGG CCGTATTCAT
4321 CCACAATTTT AAGCGGAAGG GGGGGATTGG GGGGTACAGT GCGGGGGAGC GGATCGTGGA
4381 CATCATCGCG ACCGACATCC AGACTAAGGA GCTGCAAAAG CAGATTACCA AGATTAGAA
4441 TTTCCGGGTC TACTACAGGG ACAGCAGAAA TCCCCTCTGG AAAGGCCAG CGAAGCTCCT
4501 CTGGAAGGGT GAGGGGGCAG TAGTGATCCA GGATAATAGC GACATCAAGG TGGTGCCAG
4561 AAGAAAGGCG AAGATCATTA GGGATTATGG CAAACAGATG GCGGGTGATG ATTGCGTGGC
4621 GAGCAGACAG GATGAGGATT AG

```

SEQ. I.D. NO. 13 - pSYNGP3 - codon optimised HIV-1 gagpol with leader sequence from the major splice donor

```

1 GTGAGTACGC CAAAAATTTT GACTAGCGGA GGCTAGAAGG AGAGAGATGG GCGCCCGCGC
61 CAGCGTGCTG TCGGGCGGCG AGCTGGACCG CTGGGAGAAG ATCCGCCTGC GCCCGGCGCG
121 CAAAAAGAAG TACAAGCTGA AGCACATCGT GTGGGCCAGC CGCGAACTGG AGCGCTTCGC
181 CGTGAACCCC GGGCTCCTGG AGACCAGCGA GGGGTGCCGC CAGATCCTCG GCCAACTGCA
241 GCCCAGCCTG CAAACCGGCA GCGAGGAGCT GCGCAGCCTG TACAACACCG TGGCCACGCT
301 GTACTGCGTC CACCAGCGCA TCGAAATCAA GGATACGAAA GAGGCCCTGG ATAAAATCGA
361 AGAGGAACAG AATAAGAGCA AAAAGAAGC CCAACAGGCC GCCGCGGACA CCGGACACAG
421 CAACCAGGTC AGCCAGAACT ACCCATCGT CAGAAACATC CAGGGGCGAGA TGGTGCACCA
481 GGCCATCTCC CCCCACGCGC TGAACGCCTG GGTGAAGGTG GTGGAAGAGA AGGCTTTTAG
541 CCCGGAGGTG ATACCCATGT TCTCAGCCCT GTCAGAGGGA GCCACCCCCC AAGATCTGAA
601 CACCATGCTC AACACAGTGG GGGGACACCA GGCCGCCATG CAGATGCTGA AGGAGACCAT
661 CAATGAGGAG GCTGCCGAAT GGGATCGTGT GCATCCGGTG CACGCAGGGC CCATCGCACC
721 GGCCAGATG CGTGAGCCAC GGGGCTCAGA CATCGCCGGA ACGACTAGTA CCCTTCAGGA
781 ACAGATCGGC TGGATGACCA ACAACCCACC CATCCCGGTG GGAGAAATCT CCAACGCTG
841 GATCATCCTG GGCCTGAACA AGATCGTGCG CATGTATAGC CCTACCAGCA TCCTGGACAT
901 CCGCCAAGGC CCGAAGGAAC CTTTTCGCGA CTACGTGGAC CGGTTCTACA AAACGCTCCG
961 CGCCGAGCAG GCTAGCCAGG AGGTGAAGAA CTGGATGACC GAAACCCTGC TGGTCCAGAA
1021 CGCGAACCCG GACTGCAAGA CGATCTGAA GGCCCTGGGC CCAGCGGCTA CCCTAGAGGA
1081 AATGATGACC GCCTGTCAGG GAGTGGGCGG ACCCGGCCAC AAGGCACGCG TCCTGGCTGA
1141 GGCCATGAGC CAGGTGACCA ACTCCGCTAC CATCATGATG CAGCGCGGCA ACTTTTCGAA
1201 CCAACGCAAG ATCGTCAAGT GCTTCAACTG TGGCAAAGAA GGGCACACAG CCCGCAACTG
1261 CAGGGCCCCCT AGGAAAAAGG GCTGTTGGAA ATGTGGAAAG GAAGGACACC AAATGAAAGA
1321 TTGTACTGAG AGACAGGCTA ATTTTTTAGG GAAGATCTGG CCTTCCACCA AGGGAAGGCC
1381 AGGGAATTTT CTTAGAGCA GACCAGAGCC AACAGCCCCA CCAGAAGAGA GCTTCAGGTT
1441 TGGGGAAGAG ACAACAATC CCTCTCAGAA CAGGAGCCG ATAGACAAG AACTGTATCC
1501 TTTAGCTTCC CTCAGATCAC TCTTTGGCAG CGACCCCTCG TCACAATAAA GATAGGGGGG
1561 CAGCTCAAGG AGGCTCTCCT GGACACCGGA GCAGACGACA CCGTGCTGGA GGAGATGTCG
1621 TTGCCAGGCC GCTGGAAGCC GAAGATGATC GGGGGAATCG GCGGTTTCAT CAAGGTGCGC
1681 CAGTATGACC AGATCCTCAT CGAAATCTGC GGCCACAAGG CTATCGGTAC CGTGCTGGTG
1741 GGCCCCACAC CCGTCAACAT CATCGGACGC AACCTGTTGA CGCAGATCCG TTGACGCTG
1801 AACTTCCCCA TTAGCCCTAT CGAGACGGTA CCGGTGAAGC TGAAGCCCGG TGAGGACGGC
1861 CCGAAGGTCA AGCAATGGCC ATTGACAGAG GAGAAGATCA AGGCACTGGT GGAGATTTGC
1921 ACAGAGATGG AAAAGGAAGG GAAAATCTCC AAGATTGGGC CTGAGAACCC GTACAACACG
1981 CCGGTGTTCC CAATCAAGAA GAAGGACTCG ACGAAATGGC GCAAGCTGGT GGACTTCCGC
2041 GAGCTGAACA AGCGCACGCA AGACTTCTGG GAGGTTGAGC TGGGCATCCC GCACCCCGCA
2101 GGGCTGAAGA AGAAGAAATC CGTGACCGTA CTGGATGTGG GTGATGCCTA CTCTCCGTT
2161 CCCCTGGACG AAGACTTCAG GAGTACACT CGCTTCACAA TCCCTTCGAT CAACAACGAG
2221 ACACCGGGGA TTCGATATCA GTACAACGTG CTGCCCCAGG GCTGGAAAGG CTCTCCCGCA

```


2281 ATCTTCCAGA GTAGCATGAC CAAAATCCTG GAGCCTTTCC GCAAACAGAA CCCCAGACATC
 2341 GTCATCTATC AGTACATGGA TGACTTGATC GTGGGCTCTG ATCTAGAGAT AGGGCAGCAC
 2401 CGCACCAAGA TCGAGGAGCT GCGCCAGCAC CTGTTGAGGT GGGGACTGAC CACACCCGAC
 2461 AAGAAGCACC AGAAGGAGCC TCCCTTCCTC TGGATGGGTT ACGAGCTGCA CCTTGACAAA
 2521 TGGACCGTGC AGCCTATCGT GCTGCCAGAG AAAGACAGCT GGACTGTCAA CGACATACAG
 2581 AAGCTGGTGG GGAAGTTGAA CTGGGCAGT CAGATTTACC CAGGGATTAA GGTGAGGCAG
 2641 CTGTGCAAAC TCCTCCGCGG AACCAAGGCA CTCACAGAGG TGATCCCCCT AACCAGGAGG
 2701 GCCGAGCTCG AACTGGCAGA AAACCGAGAG ATCCTAAAGG AGCCCGTGCA CGGCGTGTAC
 2761 TATGACCCCT CCAAGGACCT GATCGCCGAG ATCCAGAAGC AGGGGCAAGG CCAGTGGACC
 2821 TATCAGATTT ACCAGGAGCC CTTCAAGAAC CTGAAGACCG GCAAGTACGC CCGGATGAGG
 2881 GGTGCCCCA CTAACGACGT CAAGCAGCTG ACCGAGGCCG TGCAGAAGAT CACCACCGAA
 2941 AGCATCGTGA TCTGGGGAAA GACTCCTAAG TTCAAGCTGC CCATCCAGAA GGAACCTGG
 3001 GAAACCTGGT GGACAGAGTA TTTGGCAGGC CCTGGATTG CTGAGTGGGA GTTCGTCAAC
 3061 ACCCCTCCCC TGGTGAAGCT GTGGTACCAG CTGGAGAAGG AGCCCATAGT GGGCGCCGAA
 3121 ACCTTCTACG TGGATGGGGC CGCTAACAGG GAGACTAAGC TGGGCAAAGC CGGATACGTC
 3181 ACTAACCGGG GCAGACAGAA GGTTGTCACC CTCACTGACA CCACCAACCA GAAGACTGAG
 3241 CTGCAGGCCA TTTACCTCGC TTTGCAGGAC TCGGGCCTGG AGGTGAACAT CGTGACAGAC
 3301 TCTCAGTATG CCTGGGCAT CATTCAAGCC CAGCCAGACC AGAGTGAGTC CGAGCTGGTC
 3361 AATCAGATCA TCGAGCAGCT GATCAAGAAG GAAAAGGTCT ATCTGGCCTG GGTACCCGCC
 3421 CACAAAGGCA TTGGCGGCAA TGAGCAGGTC GACAAGCTGG TCTCGGCTGG CATCAGGAAG
 3481 GTGCTATTCC TGGATGGCAT CGACAAGGCC CAGGACGAGC ACGAGAAATA CCACAGCAAC
 3541 TGGCGGGCCA TGGCTAGCGA CTTCAACCTG CCCCCTGTGG TGGCCAAAGA GATCGTGGCC
 3601 AGCTGTGACA AGTGTACGCT CAAGGGCGAA GCCATGCATG GCCAGGTGGA CTGTAGCCCC
 3661 GGCACTGTGC AACTCGATTG AACTCATCTG GAGGGCAAGG TTATCCTGGT AGCCGTCCAT
 3721 GTGGCCAGTG GCTACATCGA GGCCGAGGTC ATTCCCGCCG AAACAGGGCA GGAGACAGCC
 3781 TACTTCCTCC TGAAGCTGGC AGGCCGCTGG CCAGTGAAGA CCATCCATAC TGACAATGGC
 3841 AGCAATTTCA CCAGTGCTAC GGTTAAGGCC GCCTGCTGGT GGGCGGGAAT CAAGCAGGAG
 3901 TTCGGGATCC CCTACAATCC CCAGAGTCAG GGCGTCGTCG AGTCTATGAA TAAGGAGTTA
 3961 AAGAAGATTA TCGGCCAGGT CAGAGATCAG GCTGAGCATC TCAAGACCCG GGTCCAAATG
 4021 GCGGTATTCA TCCACAATTT CAAGCGGAAG GGGGGGATTG GGGGGTACAG TCGGGGGGAG
 4081 CGGATCGTGG ACATCATCGC GACCGACATC CAGACTAAGG AGCTGCAAAA GCAGATTACC
 4141 AAGATTCAGA ATTTCCGGGT CTACTACAGG GACAGCAGAA ATCCCCTCTG GAAAGGCCCA
 4201 GCGAAGCTCC TCTGGAAGGG TGAGGGGGCA GTAGTGATCC AGGATAATAG CGACATCAAG
 4261 GTGGTGCCCA GAAGAAAGGC GAAGATCATT AGGGATTATG GCAAACAGAT GGCGGGTGAT
 4321 GATTGCGTGG CGAGCAGACA GGATGAGGAT TAG

SEQ. I.D. NO. 14 – pSYNGP4 – codon optimised HIV-1 gagpol with 20 bp of the leader sequence of HIV-1, upstream of the start codon of ATG.

1 CGGAGGCTAG AAGGAGAGAG ATGGGCGCCC GCGCCAGCGT GCTGTGCGGC GGCGAGCTGG
 61 ACCGCTGGGA GAAGATCCGC CTGCGCCCCG GCGGCCAAAA GAAGTACAAG CTGAAGCACA
 121 TCGTGTGGGC CAGCCGCGAA CTGGAGCGCT TCGCCGTGAA CCCCAGGCTC CTGGAGACCA
 181 GCGAGGGGTG CCGCCAGATC CTCGGCCAAC TGCAGCCAG CCTGCAAACC GGCAGCGAGG
 241 AGCTGCGCAG CCTGTACAAC ACCGTGGCCA CGCTGTACTG CGTCCACCAG CGCATCGAAA
 301 TCAAGGATAC GAAAGAGGCC CTGGATAAAA TCGAAGAGGA ACAGAATAAG AGCAAAAAAG
 361 AGGCCCAACA GGCCGCGCGG GACACCGGAC ACAGCAACCA GGTGAGCCAG AACTACCCCA
 421 TCGTGCAGAA CATCCAGGGG CAGATGGTGC ACCAGGCCAT CTCCCCCGC ACGCTGAACG
 481 CCTGGGTGAA GGTGGTGGAA GAGAAGGCTT TTAGCCCGGA GGTGATACCC ATGTTCTCAG
 541 CCCTGTGAGA GGGAGCCACC CCCCAGATC TGAACACCAT GCTCAACACA GTGGGGGGAC
 601 ACCAGGCCCG CATGCAGATG CTGAAGGAGA CCATCAATGA GGAGGCTGCC GAATGGGATC
 661 GTGTGCATCC GGTGCACGCA GGGCCCATCG CACCGGGCCA GATGCGTGAG CCACGGGGCT
 721 CAGACATCGC CGGAACGACT AGTACCCCTT AGGAACAGAT CGGCTGGATG ACCAACAACC
 781 CACCCATCCC GGTGGGAGAA ATCTACAAAC GCTGGATCAT CCTGGGCTG AACAAGATCG
 841 TGCGCATGTA TAGCCCTACC AGCATCCTGG ACATCCGCCA AGGCCGAAG GAACCCCTTC
 901 GCGACTACGT GGACCGGTTT TACAAAACGC TCCGCGCCGA GCAGGCTAGC CAGGAGGTGA
 961 AGAATCTGGT GACCGAAACC CTGTCTGGTC AGAACGCGAA CCCGGAAGTC AAGACGATCC
 1021 TGAAGGCCCT GGGCCAGCG GGTACCTAG AGGAAATGAT GACCGCTGTG GACCGAGTGG
 1081 GCGGACCCCG CCACAAGGCA CGCGTCTGG CTGAGGCCAT GAGCCAGGTG ACCAACTCCG
 1141 CTACCATCAT GATGCAGCGC GGCAACTTTC GGAACCAACG CAAGATCGTC AAGTGCTTCA
 1201 ACTGTGGCAA AGAAGGGCAC ACAGCCCGCA ACTGCAGGGC CCCTAGGAAA AAGGGCTGTT
 1261 GGAAATGTGG AAAGGAAGGA CACCAATGTA AAGATTGTAC TGAGAGACAG GCTAATTTTT
 1321 TAGGAAGAT CTGGCCTTCC CACAAGGGAA GGCAGGGGAA TTTTCTTCAG AGCAGACCAG
 1381 AGCCAAGACG CCCACCAGAA GAGAGCTTCA GGTGTTGGGA AGAGACAACA ACTCCCTCTC
 1441 AGAAGCAGGA GCCGATAGAC AAGGAAGTGT ATCCTTTAGC TTCCCTCAGA TCACCTTTTG

1501 GCAGCGACCC CTCGTCACAA TAAAGATAGG GGGGCAGCTC AAGGAGGCTC TCCTGGACAC
1561 CGGAGCAGAC GACACCGTGC TGGAGGAGAT GTCGTTGCCA GGCCGCTGGA AGCCGAAGAT
1621 GATCGGGGGA ATCGGCGGTT TCATCAAGGT GCGCCAGTAT GACCAGATCC TCATCGAAAT
1681 CTGCGGCCAC AAGGCTATCG GTACCGTGCT GGTGGGCCCC ACACCCGTCA ACATCATCGG
1741 ACGCAACCTG TTGACGCAGA TCGGTTGCAA GCTGAACTTC CCCATTAGCC CTATCGAGAC
1801 GGTACCGGTG AAGCTGAAGC CCGGGATGGA CGCCCCGAAG GTCAAGCAAT GGCCATTGAC
1861 AGAGGAGAAG ATCAAGGCAC TGGTGGAGAT TTGCACAGAG ATGGAAAAGG AAGGGAAAAT
1921 CTCCAAGATT GGGCCTGAGA ACCCGTACAA CACGCCGGTG TTCGCAATCA AGAAGAAGGA
1981 CTCGACGAAA TGGCGCAAGC TGGTGGACTT CCGCGAGCTG AACAAAGCGCA CGCAAGACTT
2041 CTGGGAGGTT CAGCTGGGCA TCCCGCACCC CGCAGGGCTG AAGAAGAAGA AATCCGTGAC
2101 CGTACTGGAT GTGGGTGATG CCTACTCTC CGTTCCCCTG GACGAAGACT TCAGGAAGTA
2161 CACTGCCTTC ACAATCCCTT CGATCAACAA CGAGACACCG GGGATTCTGAT ATCAGTACAA
2221 CGTGCTGCCC CAGGGCTGGA AAGGCTCTCC CGCAATCTTC CAGAGTAGCA TGACCAAAAT
2281 CCTGGAGCCT TTCCGCAAAC AGAACCCCGA CATCGTCATC TATCAGTACA TGGATGACTT
2341 GTACGTGGGC TCTGATCTAG AGATAGGGCA GCACCCGACC AAGATCGAGG AGCTGCGCCA
2401 GCACCTGTTG AGGTGGGGAC TGACCACACC CGACAAGAAG CACCAGAAGG AGCCTCCCTT
2461 CCTCTGGATG GGTACGAGC TGCACCCTGA CAAATGGACC GTGCAGCCTA TCGTGCTGCC
2521 AGAGAAAGAC AGCTGGACTG TCAACGACAT ACAGAAGCTG GTGGGGAAGT TGAAGTGGGC
2581 CAGTCAGATT TACCCAGGGA TTAAGGTGAG GCAGCTGTGC AAATCCTCTC GCGGAACCAA
2641 GGCACCTACA GAGGTGATCC CCCTAACCGA GGAGGCCGAG CTCGAACTGG CAGAAAACCG
2701 AGAGATCCTA AAGGAGCCCG TGCACGGCGT GTACTATGAC CCCTCCAAGG ACCTGATCGC
2761 CGAGATCCAG AAGCAGGGGC AAGGCCAGTG GACCTATCAG ATTTACCAGG AGCCTTCAA
2821 GAACCTGAAG ACCGGCAAGT ACGCCCGGAT GAGGGGTGCC CACACTAACG ACGTCAAGCA
2881 GCTGACCGAG GCCGTGCAGA AGATCACCAC CGAAAGCATC GTGATCTGGG GAAAGACTCC
2941 TAAGTTCAAG CTGCCCATCC AGAAGGAAAC CTGGGAAACC TGGTGGACAG AGTATTGGCA
3001 GGCCACCTGG ATTCTGAGT GGGAGTTCGT CAACACCCCT CCCCTGGTGA AGCTGTGGTA
3061 CCAGCTGGAG AAGGAGCCCA TAGTGGGCGC CGAAACCTTC TACGTGGATG GGGCCGCTAA
3121 CAGGGAGACT AAGCTGGGCA AAGCCGGATA CGTCACTAAC CGGGGCAGAC AGAAGGTTGT
3181 CAGCCTCACT GACACCACCA ACCAGGAGAC TGAGCTGCAG GCCATTTACC TCGCTTTGCA
3241 GGACTCGGGC CTGGAGGTGA ACATCGTGAC AGACTCTCAG TATGCCCTGG GCATCATTCA
3301 AGCCCAGCCA GACCAGAGTG AGTCCGAGCT GGTCAATCAG ATCATCGAGC AGCTGATCAA
3361 GAAGGAAAAG GTCTATCTGG CCTGGGTACC CGCCACAA GGCATTGGCG GCAATGAGCA
3421 GGTCGACAAG CTGGTCTCGG CTGGCATCAG GAAGGTGCTA TTCCTGGATG GCATCGACAA
3481 GGCCAGGAC GAGCACGAGA AATACCACAG CAACTGGCGG GCCATGGCTA GCGACTTCAA
3541 CCTGCCCCCT GTGGTGGCCA AAGAGATCGT GGCCAGCTGT GACAAGTGTC AGCTCAAGGG
3601 CGAAGCCATG CATGGCCAGG TGGACTGTAG CCCCGGCATC TGGCAACTCG ATTGCACCCA
3661 TCTGGAGGGC AAGGTTATCC TGGTAGCCGT CCATGTGGCC AGTGGCTACA TCGAGGCCGA
3721 GGTCATTCCC GCCGAAACAG GGCAGGAGAC AGCCTACTTC CTCCTGAAGC TGGCAGGCCG
3781 GTGGCCAGTG AAGACCATCC ATACTGACAA TGGCAGCAAT TTCACCAGTG CTACGGTTAA
3841 GGCCGCCTGC TGGTGGCGG GAATCAAGCA GGAGTTCGGG ATCCCCTACA ATCCCCAGAG
3901 TCAGGGCGTC GTCGAGTCTA TGAATAAGGA GTTAAAGAAG ATTATCGGCC AGGTCAGAGA
3961 TCAGGCTGAG CATCTCAAGA CCGCGGTCCA AATGGCGGTA TTCATCCACA ATTTCAAGCG
4021 GAAGGGGGGG ATTGGGGGGT ACAGTGCAGG GGAGCGGATC GTGGACATCA TCGCGACCGA
4081 CATCCAGACT AAGGAGCTGC AAAAGCAGAT TACCAAGATT CAGAATTTCC GGGTCTACTA
4141 CAGGGACAGC AGAAATCCCC TCTGGAAAGG CCCAGCGAAG CTCCTCTGGA AGGGTGAGGG
4201 GGCAGTAGTG ATCCAGGATA ATAGCGACAT CAAGGTGGTG CCCAGAAGAA AGGCGAAGAT
4261 CATTAGGGAT TATGGCAAAC AGATGGCGGG TGATGATTGC GTGGCGAGCA GACAGGATGA
4321 GGATTAG